

Round 3 results — Run fusionofruns submitted from IRLabKU

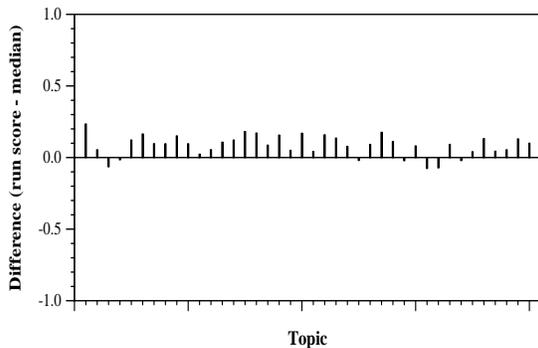
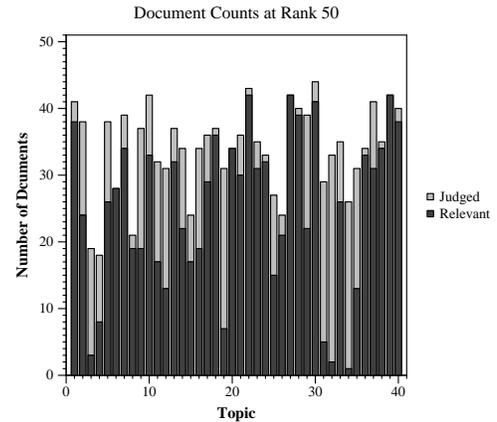
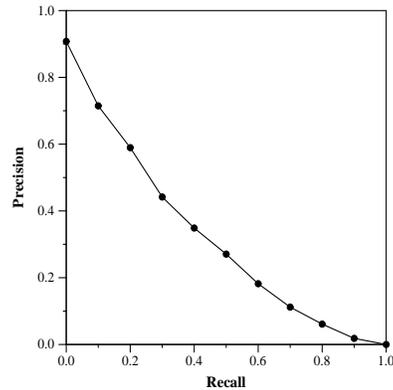
Run Description

- We create 3 indexes: 1) Based on title+abstract, 2) Based on full-text, and 3) Based on title+abstract+paragraph (a document is split into para1, para2, ..., paraK, and we create K+1 documents of the form title+abstract, title+abstract+para1, title+abstract+para2. - We create 4 different types of queries: 1) Query + named entities(question), 2) Query + named entities(question) + named entities(narrative), 3) Question + named entities(query), and 4) Question + named entities(query) + named entities(narrative) - For each query type we create 4 runs (default BM25 + default RM3): 1) We search in index of title+abstract, 2) We search in index of full-text, 3) We search in index of title+abstract+paragraph (where we only keep the first occurrence of a document and delete the following duplicates), and 4) We search in index of title+abstract+paragraph (where we don't remove the duplicates) Finally, based on each of these runs, we use reciprocal rank fusion to get 1 run.

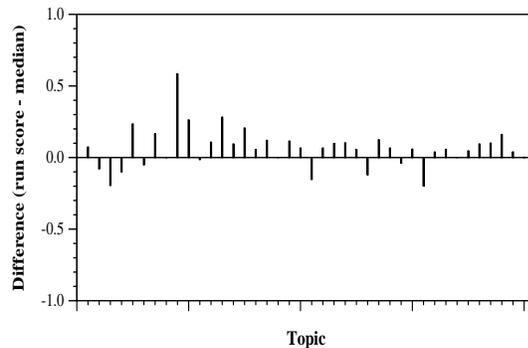
Summary Statistics	
Run ID	fusionofruns
Topic type	automatic
Contributed to judgment sets?	yes

Overall measures	
Number of topics	40
Total number retrieved	39916
Total relevant	4698
Total relevant retrieved	3295
MAP	0.3084
Mean Bpref	0.5487
Mean NDCG@10	0.6564
Mean RBP(p=0.5)	0.6692 +0.0581

Document Level Averages	
	Precision
At 5 docs	0.7350
At 10 docs	0.7400
At 15 docs	0.6850
At 20 docs	0.6587
At 30 docs	0.5975
R-Precision	
Exact	0.3483

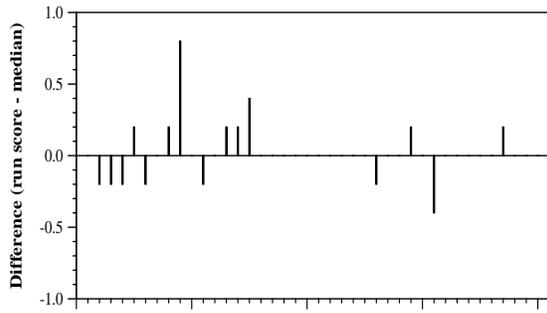


Per-topic difference from median bpref for all Round 3 runs

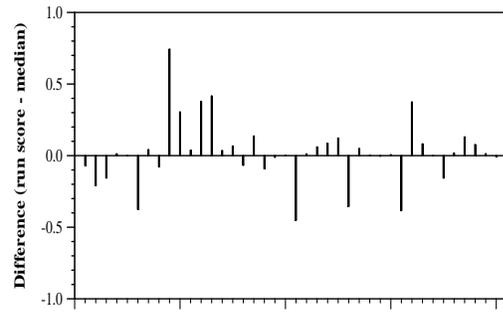


Per-topic difference from median NDCG@10 for all Round 3 runs

Round 3 results — Run fusionofruns submitted from IRLabKU



Per-topic difference from median P@5 for all Round 3 runs



Per-topic difference from median RBP(p=0.5) for all Round 3 runs